(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.[7]
G10L 11/00

(11) 공개번호 10-2004-0004420
(43) 공개일자 2004년01월13일

(21) 출원번호 10-2003-7005883 (분할)
(22) 출원일자 2003년04월28일
(62) 원출원 특허 10-1996-0701839
원출원일자 : 1996년04월09일 심사청구일자 2000년08월01일
번역문 제출일자 2003년04월28일
(86) 국제출원번호 PCT/US1995/009830 (87) 국제공개번호 WO 1996/05592
(86) 국제출원출원일자 1995년08월01일 (87) 국제공개일자 1996년02월22일

(30) 우선권주장 08/288,413 1994년08월10일 미국(US)

(71) 출원인 콸콤 인코포레이티드
미국 캘리포니아 샌디에고 모어하우스 드라이브5775 (우 92121-1714)

(72) 발명자 데자코,앤드류,피.
미국92126캘리포니아샌디에고플랜더스커브10424

가드너,윌리암,알.
미국92130캘리포니아샌디에고엑스베리코오트4650

(74) 대리인 남상선

심사청구 : 있음

(54) 가변율 보코더의 인코딩 속도를 선택하기 위한 방법 및 장치

요약

본 발명은 낮은 에너지 무성음이 배경잡음으로 코딩될 확률을 감소시키는 방법을 제공한다. 인코딩 속도는 상기 입력 신호를 디지탈 서브대역 필터들(4, 6)을 사용하여 서브대역으로 나누고, 서브대역 속도 결정 구성요소(12, 14)에서 상기 대역들에 있는 상기 에너지를 임계값들의 세트에 비교하여, 상기 비교들을 인코딩 속도 선택기(16)에서 검사함으로써 결정된다. 상기 방법에 의해, 무성음은 배경 잡음과 구별될 수 있다. 본 발명은 또한, 상기 입력 신호의 신호 대 잡음 비를 사용하여 상기 임계값 레벨을 선택하는 수단을 제공하며, 음악을 배경 잡음으로부터 구별하기 위해 상기 입력 신호의 주기성을 검사함으로써 가변율 보코더를 통과하는 음악을 코딩하는 방법을 제공한다.

대표도

도 1

명세서

도면의 간단한 설명

제 1도는 본 발명의 블록도이다.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 보코더에 관한 것으로, 특히 가변율 보코더에서 음성 인코딩 속도를 결정하기 위한 새롭고 개선된 방법에 관한 것이다.

가변율 음성 표현 시스템은 일반적으로 인코딩을 시작하기 전에 일정한 속도 결정 알고리즘의 형태를 사용한다. 상기 속도 결정 알고리즘은 음성이 존재하는 오디오 신호의 세그먼트에 대해서는 높은 비트 속도 인코딩 구성을, 무성 세크먼트에 대해서는 저속 인코딩 구성을 할당한다. 이러한 방법을 사용하여 재구성된 음성 품질을 높이면서도 평균적으로 낮은 비트 속도를 달성할 수 있다. 그러므로 가변율 음성 코더를 효과적으로 작동하려면, 다양한 잡음 환경에서 무성(unvoiced speech)으로부터 음성을 구별할 수 있는 강한 속도 결정 알고리즘이 필요하다.

이와같은 가변율 음성 압축 시스템 또는 가변율 보코더는 '가변율 보코더'라는 발명의 명칭으로 1991년 6월 11일에 출원하고 현재 계류중인 미합중국 특허출원 제 07/713,661호에 개시되어 있다. 이러한 가변율 보코더의 특수한 실시에 있어서, 입력 음성은 결정된 음성활성도의 레벨에 상응하여 여러 가지 속도 중 하나의 속도로 코드 여기 선형 예측 코딩(CELP) 방식을 사용하여 인코딩한다. 음성 활성도의 레벨은 음성에 더하여 배경 잡음을 포함할 수 있는 입력 오디오 샘플의 에너지로부터 결정된다. 보코더에 배경 잡음의 가변 레벨에 걸쳐 고품질 음성 인코딩을 제공하기 위한 보코더를 위해, 상기 속도 결정 알고리즘상에서 배경 잡음의 효과를 보상하는 임계값 기술을 적절히 사용하고 있다.

보코더는 일반적으로 전송을 위한 디지탈 형태로 전환되는 아날로그 오디오 신호의 디지탈 신호 압축을 제공하기 위하여 셀룰라 전화 또는 개인 통신 장치와 같은 통신장치에 사용된다. 셀룰라 전화 또는 개인 통신 장치가 사용될 수도 있는 이동 환경에 있어서, 높은 레벨의 배경 잡음 에너지 때문에 속도 결정 알고리즘은 신호 에너지를 사용하여 배경 잡음 침묵(silence)으로부터 저에너지 무성음을 구별하기가 어렵게 된다. 그러므로 무성음은 자주 저비트 속도에서 인코딩되며 's', 'x', 'ch', 'sh', 't'등과 같은 자음으로 재구성된 음성에서 상실되므로 음성 품질이 저하된다.

배경 잡음 에너지만을 속도 결정의 기초로 하는 보코더는 임계치 설정에 있어서 배경 잡음에 관련된 신호 세기를 고려하지 않는다. 배경 잡음에만을 그 임계 레벨의 기초로 하는 보코더는 배경 잡음이 증가하는 경우에, 임계 레벨을 서로 압축하려는 경향이 있다. 그러나, 신호 레벨이 변하지 않을 경우에 이러한 방식으로 임계치를 설정하는 것은 옳은 접근방법이지만, 신호 레벨이 배경 잡음 레벨과 함께 증가하는 경우에는 임계치를 압축하는 것은 최선의 방법이 아니다. 신호 세기를 고려한 임계 레벨를 설정하기 위한 다른 방법이 가변율 보코더에 필요하다.

남아있는 마지막 문제는 배경 잡음 에너지에 근거한 속도 결정 보코더를 통해 음악이 연주되는 동안에 일어난다. 사람들이 말할 때, 임계 레벨을 적당한 배경 잡음 레벨로 재설정하기 위해 숨을 잠시 멈추어야만 한다. 그러나, 음악이 계속되는 상태에서 일어나는 것과 같이 보코더를 통해 음악을 전송함에 있어서, 어떠한 일시 정지도 발생하지 않으며 임계 레벨은 음악이 완전속도(full rate) 이하의 속도로 코드화되기 시작할 때 까지 연속해서 증가할 것이다. 이와같은 상태에서 가변율 보코더는 음악을 배경 잡음으로 혼동하게 된다.

발명이 이루고자 하는 기술적 과제

본 발명은 가변율 보코더에서 인코딩 속도를 결정하기 위한 새롭게 개선된 방법 및 장치이다. 본 발명의 제1목적은 저에너지 무성음이 배경잡음으로 코딩될 확률을 감소시키기 위한 방법을 제공하는 것이다. 본 발명에 있어서, 입력 신호는 고주파 성분 및 저주파 성분으로 필터링된다. 입력신호의 필터링된 성분들은 음성의 존재를 검출하기 위하여 개별적으로 분석된다. 무성음은 고주파 성분을 가지기 때문에 고주파 대역에 관련한 그 신호 세기는 전체 주파수 대역에 걸쳐 배경 잡음과 비교하는 것보다 상기 대역에서 배경잡음으로부터 보다 더 구별될 수 있다.

본 발명의 제2목적은 배경 잡음 에너지뿐만 아니라 신호 에너지를 고려한 임계치 설정 수단을 제공하는 것이다. 본 발명에 있어서, 음성 검출 임계 설정은 입력 신호의 신호 대 잡음비(SNR)의 평가에 의거한다. 전형적인 실시예에 있어서, 신호 에너지는 음성 활성 시간동안 최대 신호 에너지로 평가되며 배경 잡음 에너지는 침묵 시간동안 최소 신호 에너지로서 평가된다.

본 발명의 제3목적은 가변율 보코더를 통과하는 음악을 코딩하기 위한 방법을 제공하는 것이다. 전형적인 실시예에 있어서, 속도 선택 장치는 임계 레벨이 증가하는 다수의 연속 프레임을 검출하고 상기 프레임 수 전반에 걸쳐서 주기

성을 체크한다. 입력 신호가 주기성이 있으면 이것은 음악의 존재를 나타낸다. 음악의 존재가 검출되면 임계치는 신호가 완전속도로 코딩되어지는 레벨에서 설정된다.

발명의 구성 및 작용

본 발명의 특징, 목적 및 이점들은 도면을 참조하여 상세한 설명으로부터 보다 명백하게 될 것이다.

제1도에서, 입력 신호 S(n)는 서브대역 에너지 계산 엘리먼트(4) 및 서브대역 에너지 계산 엘리먼트(6)에 제공된다. 입력 신호 S(n)는 오디오 신호 및 배경 잡음으로 이루어진다. 오디오 신호는 전형적으로 음성이지만 음악일수도 있다. 전형적인 실시예에서, S(n)은 20ms 프레임의 160 샘플에 각각 제공된다. 전형적인 실시예에 있어서, 입력 신호 S(n)은 대략 사람 음성 신호의 대역폭인 0 kHz 내지 4 kHz의 주파수 성분을 갖는다.

전형적인 실시예에 있어서, 4 kHz 입력 신호 S(n)는 두 개의 분리된 서브대역으로 필터링된다. 두 개의 분리된 서브대역은 각각 0 kHz 내지 2 kHz와 2 kHz 내지 4 kHz 사이에 놓인다. 전형적인 실시예에 있어서, 입력 신호는 서브대역 필터에 의해 서브대역들로 분할될 수도 있으며, 그 설계는 당해분야에서 공지이고 ' 주파수 선택 적응 필터링'이란 발명의 명칭으로 1994년 2월 1일에 출원한 미합중국 특허출원 제 08/189,819호에 상세되어 있다.

서브대역 필터의 임펄스 응답은 저대역통과 필터에 대해서는 $h_L(n)$, 고대역통과 필터에 대해서는 $h_H(n)$으로 규정된다. 신호의 서브대역 성분의 에너지는 당해분야에서 공지인 바와같이, 상기 서브대역 필터 출력 샘플들의 제곱을 합함으로써 $R_L(0)$ 및 $R_H(0)$을 제공하도록 간단히 계산될 수 있다.

바람직한 실시예에 있어서, 입력 신호 S(n)가 서브대역 에너지 계산 엘리먼트(4)에 제공되는 경우, 입력 프레임의 저주파 성분의 에너지값,$R_L(0)$,은 다음과 같이 계산된다.

$$R_L(0) = R_S(0) \cdot R_{hL}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{hL}(i), \qquad (1)$$

여기서 L은 임펄스 응답 hL(n)을 갖는 저대역통과 필터에서 탭수이며, $R_S(i)$는 다음 방정식에 의해 주어진 입력 신호 S(n)의 자기상관 함수이다 :

$$\cdot R_S(i) = \sum_{n=1}^{N} S(n) \cdot S(n-i), \qquad \text{for } i \in [0,\ L-1] \qquad (2)$$

여기서 N은 프레임의 샘플수이며, RhL은 다음 방정식에 의해 주어진 저대역통과 필터 $h_L(n)$의 자기상관 함수이다 :

$$\begin{cases} R_{hL}(i) = \sum_{n=0}^{L-1} h_L(n) \cdot h_L(n-i). & \text{for } i \in [0, L-1] \qquad (3) \\ = 0 & else \end{cases}$$

고주파 에너지 $R_H(0)$는 서브대역 에너지 계산 엘리먼트(6)에서 유사한 방식으로 계산된다.

서브대역 필터의 자기상관 함수의 값은 계산 로드를 감소시키기 위해 시간에 앞서 계산될 수 있다. 또한, $R_S(i)$의 계산된 값의 일부는 입력 신호 S(n)의 코딩에서 다른 계산에 사용되며, 그것에 의해 본 발명의 인코딩 속도 선택 방법의 네트 계산 부담이 더 감소된다. 예를 들면, LPC 필터 탭 값의 유도는 일 세트의 입력 신호 자기상관 계수의 계산을 필요로 한다.

LPC 필터 탭 값의 계산은 당해 분야에서 공지이며 미합중국 특허출원 제 08/004,484 호에 상세하게 설명되어 있다. 10 탭 LPC 필터를 필요로 하는 방법으로 음성을 코딩하면 11부터 L-1까지 i값에 대한 $R_S(i)$의 값만이 계산될 필요가 있으며, 그것에 더하여, LPC 필터 탭 값을 계산하는데 0 내지 10까지 i값에 대한 RS(i)이 사용되기 때문에 상기 11부터 L-1까지 i값에 대한 RS(i)이 신호를 코딩하는데 사용된다. 전형적인 실시예에 있어서, 서브대역 필터는 17 탭을 갖는다. 즉 L=17이다.

서브대역 에너지 계산 엘리먼트(4)는 서브대역 속도 결정 엘리먼트(12)에 $R_L(0)$의 계산된 값을 제공하며, 서브대역 에너지 계산 엘리먼트(6)는 서브대역 속도 결정 엘리먼트(14)에 RH(0)의 계산된 값을 제공한다. 속도 결정 엘리먼트(12)는 두 개의 소정의 임계치 $T_{L1/2}$ 및 T_{Lfull} 에 대하여 $R_L(0)$값을 비교하며, 상기 비교 에 따라서 제안된 인코딩

속도 RATE $_L$ 를 할당한다. 속도 할당은 다음과 같이 이루어진다 :

RATE $_L$ = 1/8 배속 $R_L(0) \leq T_{L1/2}$ (4)

RATE $_L$ = 1/2 배속 $T_{L1/2} < R_L(0) \leq T_{Lfull}$ (5)

RATE $_L$ = 전속 $R_L(0) > T_{Lfull}$ (6)

서브대역 속도 결정 엘리먼트(14)는 유사한 방식으로 동작하며 고주파 에너지 값$R_H(0)$에 따라서 그리고 다른 세트의 임계치 $T_{H1/2}$ 및 T_{Hfull} 에 의거하여 제안 인코딩 속도 RATE $_H$ 를 선택한다. 서브대역 속도 결정 엘리먼트(12)는 인코딩 속도 선택 엘리먼트(16)에 그 제안된 인코딩 속도 RATE $_L$ 를 제공하며, 서브대역 속도 결정 엘리먼트(14)는 인코딩 속도 선택 엘리먼트(16)에 그 제안된 인코딩 속도 RATE $_H$ 를 제공한다. 전형적인 실시예에 있어서, 인코딩 속도 선택 엘리먼트(16)는 두 개의 제안 속도들 중에서 더 빠른 속도를 선택하며, 상기 더 빠른 속도를 선택된 인코딩 속도로 제공한다.

또한 서브대역 에너지 계산 엘리먼트(4)는 임계 적응 엘리먼트(8)에 저주파 에너지값 $R_L(0)$를 제공하며, 여기서 다음 입력 프레임에 대한 임계치 $T_{L1/2}$ 및 T_L full이 계산된다. 유사하게, 서브대역 에너지 계산 엘리먼트(6)는 임계 적응 엘리먼트(10)에 고주파 에너지값 $R_H(0)$를 제공하며, 여기서 다음 입력 프레임에 대한 임계 치 $T_{H1/2}$ 및 T_{Hfull} 이 계산된다.

임계 적응 엘리먼트(8)는 저주파 에너지값 $R_L(0)$을 수신하여 S(n)이 배경 잡음이나 오디오 신호중 어느 것을 포함하고 있는지를 결정한다. 전형적인 실시예 있어서, 임계 적응 엘리먼트(8)가 오디오 신호가 존재하는지를 결정하는 방법은 I번째 프레임에 대한 정규화된 자기상관 함수 NACF(i)를 검사하는 것이며, 상기 NACF(i)는 다음과 같은 방정식에 의해 주어진다 :

$$\mathbf{NACF(i)} = \max_{T} \frac{\sum_{n=0}^{N-1} e(n) \cdot e(n-T)}{\frac{1}{2}\left[\sum_{n=0}^{N-1} e^2(n) + \sum_{n=0}^{N-1} e^2(n-T)\right]'} \quad (7)$$

여기서 e(n)은 LPC 필터에 의해 입력 신호 S(n)을 필터링함으로써 생긴 포르만트(formant) 잔여 신호이다.

LPC 필터에 의해 신호를 필터링하는 설계 구조는 당해분야에서 공지이며 미합중국 특허출원 제 08/004,484호에 상세하게 설명되어 있다. 입력신호 S(n)은 포르만트의 상호작용을 제거하기 위하여 LPC 필터에 의해 필터링된다. NACF는 오디오 신호가 존재하는지를 결정하기 위하여 임계치에 대하여 비교된다. NACF가 소정의 임계치보다 더 크면, 입력 프레임은 음성 또는 음악과 같은 오디오 신호의 존재를 나타내는 주기적 특성을 갖는다는 것을 나타낸다. 음성 또는 음악은 부분적으로 주기적이 아니며 낮은 NACF 값을 나타낼 수 있지만, 배경 잡음은 전형적으로 어떤 주기성도 디스플레이하지 않으며 거의 항상 낮은 NACF 값을 나타낸다.

S(n)이 배경 잡음을 포함하는 것으로 결정되면, NACF의 값은 임계치 TH1 이하이며, $R_L(0)$ 값은 현재의 배경 잡음 평가 BGN $_L$ 의 값을 갱신하는데 사용된다. 전형적인 실시예에 있어서, TH1은 0.35 이다. $R_L(0)$은 배경 잡음 평가 BGN $_L$ 의 현재값에 대하여 비교된다. $R_L(0)$이 BGN $_L$ 이하이면, 배경 잡음 평가 BGN $_L$ 은 NACF의 값에 관계없이 $R_L(0)$와 동일하게 설정된다.

배경 잡음 평가 BGN $_L$ 은 NACF가 임계치 TH1보다 작은 경우에만 증가된다. $R_L(0)$이 BGN $_L$ 보다 크고 NACF가 TH1보다 작으면, 배경 잡음 평가 BGN $_L$ 은 $\alpha 1 \cdot$ BGN $_L$ 로 설정되며, 여기서 $\alpha 1$은 1보다 큰 수이다. 전형적인 실시예에 있어서, $\alpha 1$은 1.03과 같다. 배경 잡음 평가 BGN $_L$ 이 BGN $_{max}$ 로 설정되는 소정의 최대값 BGN $_{max}$ 에 이를 때 까지, NACF가 임계치 TH1보다 적고 $R_L(0)$이 현재의 BGN $_L$ 값보다 클 동안은 BGN $_L$ 은 계속해서 증가할 것이다.

제2 임계치 TH2를 초과하는 NACF의 값으로 나타나는 오디오 신호가 검출되면, 신호 에너지 평가 S_L 이 갱신된다. 전형적인 실시예에 있어서, TH2는 0.5로 설정된다. $R_L(0)$의 값은 현재의 저대역통과 신호 에너지 평가 S_L 에 대하여 비교된다. $R_L(0)$이 S_L 의 현재값보다 크면 S_L 은 $R_L(0)$와 동일하게 설정된다. $R_L(0)$이 S_L 의 현재값보다 작으며 NACF가 TH2보다 큰 경우에만, S_L 은 $\alpha 2 \times R_L(0)$로 설정된다. 전형적 인 실시예에서, $\alpha 2$는 0.96으로 설정된다.

임계 적응 엘리먼트(8)는 다음 방정식(8)에 따라서 신호 대 잡음비 평가를 계산한다 :

$$SNR_L = 10 \cdot \log[\frac{S_L}{BGN_L}] \qquad (8)$$

임계 적응 엘리먼트(8)는 다음 방정식(9)-(12)에 따라서 양자화된 신호 대 잡음비I $_{SNRL}$ 의 인덱스를 결정한다

$$I_{SNRL}=\text{nint}\ [\frac{SNR_L-20}{5}], \quad \text{for } 20<SNR_L<55, \qquad (9)$$

$$=0, \quad \text{for } SNR_L \leq 20, \qquad (10)$$

$$=7 \quad \text{for } SNR_L \geq 55.$$

여기서 nint는 분수값을 가장 가까운 정수로 표시하는 함수이다.

임계 적응 엘리먼트(8)는 신호 대 잡음비 인덱스 I $_{SNRL}$ 에 따라서 두 개의 스케일 인자 k $_{L1/2}$ 및 k $_{Lfull}$ 을 선택 또는 계산한다. 전형적인 스케일 값 조사표는 표1에 제공되어 있다 :

표 1

I $_{SNRL}$ K $_{L1/2}$ K $_{Lfull}$

0 7.0 9.0

1 7.0 12.6

2 8.0 17.0

3 8.6 18.5

4 8.9 19.4

5 9.4 20.9

6 11.0 25.5

7 15.8 39.

이 두 개의 값은 다음 방정식에 따라서 속도 선택을 위한 임계치를 계산하는데 사용된다 :

T $_{L1/2}$ = K $_{L1/2 \cdot BGNL}$ (11)

T $_{LFULL}$ = K $_{LFULL \cdot BGNL}$ (12)

여기서 T $_{L1/2}$ 은 저주파 1/2 속도 임계치이며 T $_{Lfull}$ 은 저주파 완전속도 임계치이다.

임계 적응 엘리먼트(8)는 속도 결정 엘리먼트(12)에 적응된 임계치 T $_{Lfull}$ 및 T $_{L1/2}$ 을 제공한다. 임계 적응 엘리먼트(8)는 유사한 방식으로 동작하며 서브대역 속도 결정 엘리먼트(14)에 임계치 T $_{Hfull}$ 및 T $_{H1/2}$ 을 제공한다.

오디오 신호 에너지 평가 S(여기서 S는 S $_{L}$ 또는 S $_{H}$ 일수 있음)의 초기값은 다음과 같이 설정된다. 신호 에너지 평가 S $_{INIT}$ 는 -18.0 dBm0로 설정되며, 여기서 3.17 dBm0는 완전 사인 파형의 신호 세기를 규정하고, 상기 사인 파형은 전형적인 실시예에서는 -8031 내지 8031 까지의 진폭범위를 갖는 디지탈 사인 파형이다. S $_{INIT}$ 는 음향 신호가 존재하는 것이 결정될 때 까지 사용된다.

음향 신호가 초기에 검출되는 방법은 임계치에 대하여 NACF값을 비교하는 것이며, NACF가 소정수의 연속 프레임에 대한 임계치를 초과하면, 음향 신호가 존재하는 것으로 결정된다. 전형적인 실시예에 있어서, NACF는 10 연속 프레임에 대한 임계치를 초과해야만 한다. 이 상태가 충족된 후에 신호 에너지 평가 S는 선행 10 프레임의 최대 신호 에너지로 설정된다.

배경 잡음 평가 BGN_L 의 초기값은 BGN_{max} 로 설정된다. 서브대역 프레임 에너지가 BGN_{max} 보다 작은 것을 수신하자마자, 배경 잡음 평가는 수신된 서브대역 에너지 레벨의 값으로 재설정되며, 배경 잡음 평가(BGN_L)의 발생은 먼저 전술한 바와같이 진행된다.

바람직한 실시예에 있어서 잔존 상태는 일련의 완전속도 음성 프레임을 뒤따라 저속 프레임이 검출될 때 활성화된다. 전형적인 실시예에 있어서, 인코딩 속도가 완전속도보다 작게 설정되고 계산된 신호 대 잡음비가 소정의 최소 SNR보다 작은 프레임을 뒤따라 완전속도로 4개의 연속적인 음성 프레임들이 인코딩 될 경우, 상기 프레임에 대한 인코딩 속도는 완전속도로 설정된다. 전형적인 실시예에 있어서 소정의 최소 SNR은 방정식 (8)에 규정된 27.5 dBas이다.

바람직한 실시예에 있어서, 잔존 프레임의 수는 신호 대 잡음비의 함수이다. 전형적인 실시예에 있어서, 잔존 프레임의 수는 다음과 같이 결정된다 :

#잔존 프레임=1 22.5 < SNR < 27.5, (13)

#잔존 프레임=2 SNR ≤ 22.5, (14)

#잔존 프레임=0 SN R ≤ 27.5. (15)

본 발명은 음악의 존재를 검출하기 위한 방법을 제공하는 것으로, 상기 음악은 전술한 바와같이 배경 잡음 측정을 재설정하게 하는 일시 정지가 없는 것이다. 음악의 존재를 검출하기 위한 방법은 음악이 콜의 시작에 존재하지 않는 것으로 추정한다. 이것에 의해 본 발명의 인코딩 속도 선택 장치는 초기 배경 잡음 에너지를 적절히 평가한다. 배경 잡음과는 달리 음악은 주기적 특성을 가지기 때문에, 본 발명은 배경 잡음으로부터 음악을 구별하기 위하여 NACF의 값을 검사한다. 본 발명의 음악 검출 방법은 다음 방정식에 따라서 평균 NACF를 계산한다 :

$$NACF_{AVE} = \frac{1}{T} \sum_{i=1}^{T} NACF(i), \qquad (16)$$

여기서 NACF(i)는 방정식(7)에서 규정되며, T는 배경 잡음의 평가된 값이 초기 배경 잡음 평가 BGN_{INT} 로부터 증가되는 연속 프레임 수이다.

배경 잡음 BGN이 소정수의 프레임 T에 대하여 증가되고 $NACF_{AVE}$ 가 소정 임계치를 초과하면, 음악이 검출되고 배경 잡음 BGN이 BGN_{INIT} 로 재설정된다. 값 T는 인코딩 속도가 완전속도 이하로 떨어지지 않도록 충분히 낮게 설정되어야 한다. 그러므로, T 값은 음향 신호 및 BGN_{INIT} 의 함수로서 설정되어야 한다.

바람직한 실시예의 전술한 설명은 당업자가 본 발명을 사용하는 것을 가능하 게 하도록 제공되어 있다. 실시예에 대한 다양한 변형은 당해업자에게 아주 명백할 것이며 여기에 기술된 일반적인 원리들은 발명적 기능을 사용하지 않고 다른 실시예에 적용될 수도 있다. 그러므로, 본 발명은 여기에 도시된 실시예에 한정되는 것은 아니며 여기에 기술된 원리 및 신규한 특징들과 일관되게 가장 광범위한 범위에 적용된다.

발명의 효과

본 발명은 상기와 같은 구성으로 인해 낮은 에너지를 가지는 무성음이 배경잡음으로 인코딩되는 것을 방지함으로써 복구된 음성의 품질을 향상시키며, 또한 상기 임계치를 정하는 경우에, 종래의 임계치 기술과는 달리 신호대 잡음비(SNR)을 고려함으로써 상기 임계치를 적정하게 결정할 수 있다. 또한, 음악은 주기적인 특성을 가지고 있다는 점을 감안하여 음악을 배경 잡음과 구별되게 인코딩할 수 있다.

(57) 청구의 범위

청구항 1.

가변율 보코더에서 인코딩 속도를 결정하기 위한 장치로서,

입력 신호를 수신하고 소정의 서브대역 에너지 계산 포맷에 상응하여 다수의 서브대역 에너지 값들을 결정하는 서브대역 에너지 계산 수단;및

상기 다수의 서브대역 에너지 값들을 수신하고 상기 다수의 서브대역 에너지 값들에 상응하여 상기 인코딩 속도를 결정하는 속도 결정 수단을 포함하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 2.
제1항에 있어서, 상기 서브대역 에너지 계산 수단은,

$$\text{서브대역 에너지} = R_S(0)\cdot R_{hbp}(0) + 2\cdot\sum_{i=1}^{L-1} R_S(i)\cdot R_{hbp}(i)$$

여기서 L은 대역 통과 필터 hbp(n)에 있는 숫자 탭이며, Rs(i)는 상기 입력 신호 S(n)의 자기상관 함수이며, R_{hbp} 는 상기 대역통과 필터 $h_{bp}(n)$의 자기상관 함수인, 상기 식에 상응하여 상기 다수의 서브대역 에너지 값을 각각 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 3.
제1항에 잇어서 상기 서브대역 에너지 계산 수단과 상기 속도 결정 수단 사이에 위치하고 있으며, 상기 서브대역 에너지 값들을 수신하고 다수의 서브대역 값들에 상응하여 인코딩 속도 임계값들의 세트를 결정하는 임계값 계산 수단을 더 포 함하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 4.
제3항에 있어서, 상기 임계값 계산 수단은 상기 다수의 서브대역 에너지 값에 상응하여 신호 대 잡음비를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치

청구항 5.
제4항에 있어서, 상기 임계값 계산 수단은 상기 신호 대 잡음 비 값에 상응하여 스케일링 값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 6.
제5항에 있어서, 임계값 계산 수단은 배경 잡음 평가를 상기 스케일링 값에 곱하여 적어도 하나의 임계값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 7.
제1항에 있어서 상기 속도 결정 수단은 상기 다수의 서브 대역 에너지 값들 중 적어도 하나를 상기 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 8.
제6항에 있어서, 상기 속도 결정 수단은 상기 다수의 서브대역 에너지 값들 중 상기 적어도 하나를 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 9.
제1항에 있어서, 상기 속도 결정 수단은 다수의 제안된 인코딩 속도들을 결정하는데, 여기서 제안된 각 인코딩 속도는 상기 다수의 서브대역 에너지 값들에 상응하며 여기서 상기 속도 결정 수단은 상기 다수의 제안된 인코딩 속도에 상응하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 10.
가변율 보코더에서 인코딩 속도를 결정하기 위한 장치로서,

입력 신호를 수신하고 상기 입력 신호에 상응하여 신호 대 잡음 비 값을 결정하는 신호 대 잡음 비 수단; 및

상기 신호 대 잡음 비 값을 수신하여 상기 신호 대 잡음 비 값에 상응하여 상기 인코딩 속도를 결정하는 속도 결정 수단을 포함하는 것을 특징으로 하는 인코딩 속도 결정 수단.

청구항 11.
가변율 보코더에서 인코딩 속도를 결정하기 위한 장치로서,

입력 신호를 수신하고 소정의 서브대역 에너지 계산 포맷에 상응하여 다수의 서브대역 에너지 값들을 결정하는 서브대역 에너지 계산기; 및

상기 다수의 서브대역 에너지 값들을 수신하여 상기 다수의 서브대역 에너지 값들에 상응하여 상기 인코딩 속도를 선택하는 속도 선택기를 포함하는 것을 특징으로 하는 인코딩 속도 결정 수단.

청구항 12.
제11항에 있어서, 상기 서브대역 에너지 계산기는,

$$\text{서브대역 에너지} = R_s(0)\cdot R_{hbp}(0) + 2\cdot\sum_{i=1}^{L-1} R_s(i)\cdot R_{hbp}(i)$$

여기서 L은 대역 통과 필터 $h_{bp}(n)$에 있는 숫자 탭이며, $R_s(i)$는 상기 입력 신호 S(n)의 자기상관 함수이며, R_{hbp}는 상기 대역통과 필터 $h_{bp}(n)$의 자기상관 함수인, 상기 식에 상응하여 상기 다수의 서브 밴드 에너지 값을 각각 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 13.
제11항에 잇어서 상기 서브대역 에너지 계산기과 상기 속도 선택기 사이에 위치하고 있으며, 상기 서브대역 에너지 값들을 수신하고 다수의 서브대역 값들에 상응하여 인코딩 속도 임계값들의 세트를 결정하는 임계값 계산기를 더 포함하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 14.
제13항에 있어서, 상기 임계값 계산기는 상기 다수의 서브대역 에너지 값들에 상응하여 신호 대 잡음 비 값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치

청구항 15.
제14항에 있어서, 상기 임계값 계산기는 상기 신호 대 잡음 비 값에 상응하여 스케일링 값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 16.
제15항에 있어서, 임계값 계산기는 배경 잡음 평가를 상기 스케일링 값에 곱하여 적어도 하나의 임계값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 17.
제11항에 있어서 상기 속도 선택기는 상기 다수의 서브 대역 에너지 값들 중 적어도 하나를 상기 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 18.
제16항에 있어서, 상기 속도 선택기는 상기 다수의 서브대역 에너지 값들 중 적어도 하나를 상기 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 19.
제11항에 있어서, 상기 속도 선택기는 다수의 제안된 인코딩 속도들을 결정하는데, 여기서 제안된 각 인코딩 속도는 상기 다수의 서브대역 에너지 값들에 상응하며 여기서 상기 속도 선택기는 상기 다수의 제안된 인코딩 속도에 상응하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 20.
가변율 보코더에서 인코딩 속도를 결정하기 위한 장치로서,

입력 신호를 수신하고 상기 입력 신호에 상응하여 신호 대 잡음 비 값을 결 정하는 신호 대 잡음 비 계산기; 및

상기 신호 대 잡음 비 값을 수신하여 상기 신호 대 잡음 비 값에 상응하여 상기 인코딩 속도를 선택하는 속도 선택기를 포함하는 것을 특징으로 하는 인코딩 속도 결정 수단.

청구항 21.
가변율 보코더에서 인코딩 속도를 결정하기 위한 방법으로서,

입력 신호를 수신하는 단계;

소정의 서브대역 에너지 계산 포맷에 상응하여 다수의 서브대역 에너지 값들을 결정하는 단계; 및

상기 다수의 서브대역 에너지 값들에 상응하여 상기 인코딩 속도를 결정하는 단계를 포함하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 22.
제21항에 있어서, 상기 다수의 서브대역 에너지 값들을 결정하는 단계는,

$$\text{서브대역 에너지} = R_s(0)\cdot R_{hbp}(0) + 2\cdot\sum_{i=1}^{L-1} R_s(i)\cdot R_{hbp}(i)$$

여기서 L은 대역 통과 필터 $h_{bp}(n)$에 있는 숫자 탭이며, $R_s(i)$는 상기 입력 신호 S(n)의 자기상관 함수이며, R_{hbp}는 상기 대역통과 필터 $h_{bp}(n)$의 자기상관 함수인, 상기 식에 상응하여 다수의 서브대역 에너지 값을 각각 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 23.
제21항에 있어서, 다수의 서브대역 에너지 값에 상응하여 인코딩 속도 임계값들을 결정하는 단계를 더 포함하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 24.
제23항에 있어서, 인코딩 속도 임계값들의 세트를 결정하는 상기 단계는 상기 다수의 서브대역 에너지 값에 상응하여 신호 대 잡음 비 값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 25.
제24항에 있어서, 인코딩 속도 임계값들을 결정하는 상기 단계는 상기 신호 대 잡음 비 값에 상응하여 스케일링 값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 26.
제25항에 있어서, 인코딩 속도 임계값들의 세트를 결정하는 상기 단계는 배경 잡음 평가를 상기 스케일링 값에 곱하여 상기 속도 임계값을 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 27.
제21항에 있어서, 상기 인코딩 속도를 결정하는 상기 단계는 상기 다수의 서브대역 에너지 값들 중에서 적어도 하나를 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 28.
제26항에 있어서, 상기 인코딩 속도를 결정하는 상기 단계는 상기 다수의 서브대역 에너지 값들 중에서 적어도 하나를 상기 적어도 하나의 임계값에 비교하여 상기 인코딩 속도를 결정하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 29.
제21항에 있어서, 상기 다수의 서브대역 에너지 값들 각각에 상응하여 제안된 인코딩 값을 발생하는 단계를 더 포함하며, 여기서 인코딩 속도를 결정하는 상기 단계는 상기 제안된 인코딩 속도들 중에서 하나를 선택하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 30.
가변율 보코더에서 인코딩 속도를 결정하는 방법으로서,

입력 신호를 수신하는 단계;

상기 입력 신호에 상응하여 신호 대 잡음 비 값을 결정하는 단계; 및

상기 신호 대 잡음 비 값에 상응하여 상기 인코딩 속도를 결정하는 단계를 포함하는 것을 특징으로 하는 인코딩 속도 결정 방법.

청구항 31.
입력 신호에 대한 인코딩 속도를 선택하는 시스템으로서,

상기 입력 신호의 각 주파수 서브대역에 대한 신호 에너지를 결정하는 서브대역 필터 서브시스템; 및

상기 입력 신호의 각 주파수 서브대역의 상기 신호 에너지에 근거하여 상기 입력 신호의 상기 인코딩 속도를 결정하는 속도 선택 서브시스템을 포함하는 것을 특징으로 하는 인코딩 속도 선택 시스템.

청구항 32.
제31항에 있어서, 상기 서브대역 필터 서브시스템은 다수의 서브대역 에너지 계산 구성요소를 포함하며, 여기서 상기 다수의 서브대역 에너지 계산 구성요소 각각은 주파수 서브대역 신호 에너지를 결정하는 것을 특징으로 하는 인코딩 속도 선택 시스템.

청구항 33.
제32항에 있어서, 상기 속도 선택 서브시스템은 다수의 임계값 적응 구성요소들를 포함하며, 상기 다수의 임계값 적응 구성요소들 각각은 상응하는 서브대역 에너지 계산 구성요소로부터의 상기 주파수 서브대역 신호 에너지를 사용하여 상기 주파수 서브대역에 오디오 신호가 존재하는지를 결정하는 것을 특징으로 하는 인코딩 속도 선택 시스템.

청구항 34.
제33항에 있어서, 각 임계값 적응 구성요소는 상기 신호 에너지와 상기 상응하는 주파수 서브대역의 잡음 평가에 근거하여 임계값을 결정하도록 구성되며, 상기 임계값은 상기 주파수 서브대역에 상기 오디오 신호가 존재하는지를 결정하도록 사용되는 것을 특징으로 하는 인코딩 속도 선택 시스템.

청구항 35.
제33항에 있어서, 상기 다수의 임계값 적응 구성요소들은 상기 입력 신호의 주파수 서브대역들의 상기 결합된 신호 에너지들에 근거하여 임계값을 결정하도록 구성되며, 상기 임계값은 상기 주파수 서브대역에 상기 오디오 신호가 존재하는지를 결정하도록 사용되는 것을 특징으로 하는 인코딩 속도 선택 시스템.

청구항 36.
입력 신호에 대한 인코딩 속도를 결정하는 장치로서,

상기 입력 신호의 각 주파수 서브대역에 오디오 신호가 존재하는지를 결정하는 오디오 신호 탐지 기기; 및

상기 입력 신호의 각 주파수 서브대역에 오디오 신호가 존재하는지에 대한 결정에 상응하여 상기 입력 신호에 대한 상기 인코딩 속도를 결정하는 인코딩 속도 선택 기기를 포함하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 37.
제36항에 있어서, 상기 오디오 신호 탐지 기기는

상기 입력 신호의 각 주파수 서브대역에 대한 신호 에너지를 결정하는 다수의 서브대역 에너지 계산 구성요소; 및

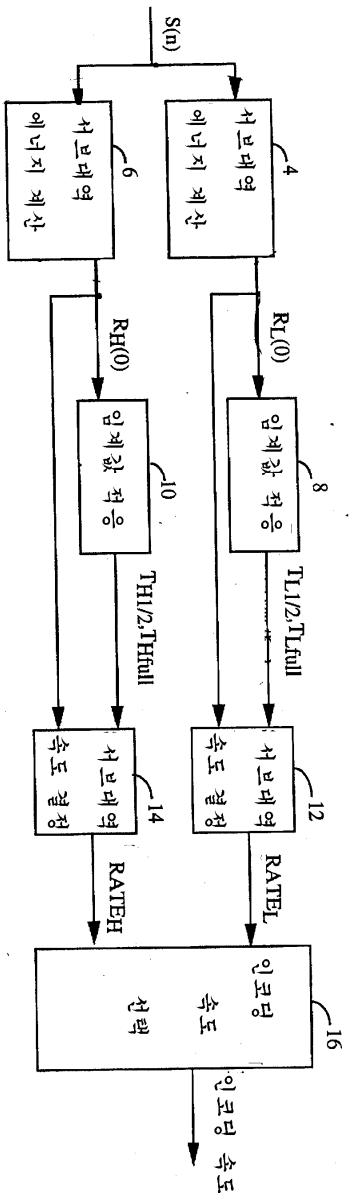
각각의 임계값 적응 구성요소가 상기 다수의 서브대역 에너지 계산 구성요소 중 하나에 통신적으로 연결되어 있으며, 할당된 주파수 서브대역의 신호 에너지를 사용하여 상기 할당된 주파수 서브대역에 오디오 신호가 존재하는지를 결정하는 단수의 임계값 적응 구성요소들을 포함하는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 38.
제37항에 있어서, 상기 인코딩 속도 선택 기기는 상기 다수의 임계값 적응 구성요소들 각각에 의해 만들어진 상기 결정에 근거하여 상기 입력 신호에 대한 상기 인코딩 속도를 선택하도록 구성되어 있는 것을 특징으로 하는 인코딩 속도 결정 장치.

청구항 39.
제38항에 있어서, 상기 다수의 임계값 적응 구성요소들 각각은 상기 할당된 주파수 서브대역의 상기 신호 에너지와 배경 잡음 평가에 근거하여 임계값을 결정하며, 상기 임계값은 상기 할당된 주파수 서브대역에 오디오 신호가 존재하는지를 결정하는데 사용되는 것을 특징으로 하는 인코딩 속도 결정 장치.

도면

도면1


S(n)
서브대역 에너지 계산
4
서브대역 에너지 계산
6
RL(0)
RH(0)
임계값 적응
8
임계값 적응
10
TL1/2,TLfull
TH1/2,THfull
서브대역 속도 결정
12
서브대역 속도 결정
14
RATEL
RATEH
인코딩 속도 선택
16
인코딩 속도